Ability sorting and the returns to college major

Peter Arcidiacono*

Department of Economics, Duke University, Box 90097, Durham, NC 27708, USA

Abstract

Large earnings and ability differences exist across majors. This paper seeks to estimate the monetary returns to particular majors as well as find the causes of the ability sorting across majors. In order to accomplish this, I estimate a dynamic model of college and major choice. Even after controlling for selection, large earnings premiums exist for certain majors. Differences in monetary returns explain little of the ability sorting across majors; virtually all ability sorting is because of preferences for particular majors in college and the workplace, with the former being larger than the latter.
© 2003 Elsevier B.V. All rights reserved.

1. Introduction

Students who choose natural science majors earn substantially more than humanities majors. In fact, economists have reported that differences in returns to majors are much larger than differences in returns to college quality. James et al. (1989, p. 252) argue that "... while sending your child to Harvard appears to be a good investment, sending him to your local state university to major in Engineering, to take lots of math, and preferably to attain a high GPA, is an even better private investment." Although a number of researchers have documented the large differences in earnings across majors (see Daymont and Andrisani, 1984; Grogger and Eide, 1995; James et al., 1989; Loury, 1997; Loury and Garman, 1995), none of the papers model the choice of major itself and we do not know whether these are actual monetary premiums or whether the observed premiums are driven by the differing abilities of individuals choosing the different majors.

---

* Tel.: +1-919-660-1816.
  E-mail address: psarcidi@econ.duke.edu (P. Arcidiacono).

The issue of selection is potentially very important as wide differences exist in the ability compositions across majors. Material covered and the jobs associated with particular majors demand different sets of skills, some of which are learned in college. The difficulty in mastering these skills may vary with ability. Turner and Bowen (1999) document the sorting that occurs across majors by SAT math and verbal scores. However, they do not link this sorting to the marketplace. In contrast, Berger (1988) relates expected earnings to major choice, yet does not discuss the vast differences in abilities across majors (ability sorting) nor does he calculate earnings premiums across majors.

This paper also links the literature on college choice to that of the choice of major. While both Fuller et al. (1982) and Brewer and Ehrenberg (1999) estimate models of college choice, no attempt has been made to integrate the college major decision with the college choice decision empirically. Yet, there are reasons to believe the decisions are linked. For example, one of Thomas Sowell's arguments against affirmative action in higher education is that individuals who choose schools where their abilities are significantly below those of their peers may be forced into easier majors (Sowell, 1972). On the other hand, should peer effects be important, individuals who attend higher quality schools may be more likely to choose lucrative majors if they are surrounded by ambitious students. Modeling the joint process makes it possible to calculate which effect is larger.

Neither the literature on college choice nor the literature on the choice of major has treated these decisions as dynamic for the individual. While transferring schools is a somewhat rare event, changing majors is not. The dynamics are important for three reasons. First, by having individuals choose over multiple time periods it is possible to separate the effect of the school environment from the effect of the workplace on the choice of college major. Second, the dynamics also allow for learning about one's abilities through grades. Those who perform worse than expected may find it more attractive to drop out or switch to a less difficult major. Finally, the dynamics make it possible to control for selection into the various stages of the model as shown by Keane and Wolpin (1997) and Cameron and Heckman (1998, 2001). In both of these papers there is no heterogeneity in school quality or in the choice of major.

Selection is the central issue in measuring the returns not only to choice of major, but to college in general. Selection can take two forms. First, even if the returns to ability are the same whether or not one has a college degree, high-ability individuals may find college to be a less costly investment. Second, the returns to college may differ across individuals; those with the highest returns may be most likely to take part in the 'treatment' of attending college (see Card, 2001; Heckman and Vytlacil, 1998; Heckman et al., 2000). This paper allows the returns to college to vary through the choice of major and through different returns to ability across majors. By explicitly modeling the educational decision process, I hope to both disentangle the heterogeneous treatment effects and control for the self-selection inherent in educational outcomes.

To accomplish this, I propose a dynamic model of college and major choice which has three periods. In the first period, individuals choose both a college and a major or choose to enter the labor force. The first period decision is made given expectations about what choices will be made in the second period. In the second period, individuals learn more about the characteristics of each of the majors as well as how they perform

in the college environment. With this new information, individuals update their decisions by changing their major, changing their college, or entering the labor force. In the third period, individuals work, receiving earnings based upon their past educational choices.[1] The model is flexible enough to capture the relationship between college quality and choice of major while allowing individuals to switch majors over time.

I find large differences in earnings premiums across majors even after controlling for selection. These differences are much larger than the returns to college quality. While math ability is an important contributor to future earnings, verbal ability is not. However, the differences in returns to math ability across schooling options explain very little of the ability sorting across majors; individuals with high math ability receive uniformly higher earnings regardless of their educational choices.

If the monetary returns do not drive the sorting, what does? High math ability individuals prefer both the subject matter and the jobs associated with the lucrative majors, with the former being larger than the latter. That is, the small differences in returns to abilities across majors are dominated by the large differences in preferences that high ability individuals have for the more lucrative fields. Further, in contrast to Sowell's argument, heterogeneous schools increase ability sorting across majors. High-quality schools actually make lucrative majors more attractive and therefore, since high-quality schools attract high-ability students, contribute to the ability sorting across majors.

Learning about one's abilities also plays a role in the choice of major. Grade residuals are approximately 50% noise, 40% ability which is transferable across majors, and 10% ability which is major specific. The new information received through grades then affects the final major choice. Those who perform worse than expected are more likely to drop out or switch to a less difficult major, while those who perform better than expected are more likely to stay in the same major or switch to a more difficult major.

The rest of the paper proceeds as follows. Section 2 provides the empirical regularities in the data. The dynamic model of college and major choice as well as the econometric techniques used to estimate the model are described in Section 3. Section 4 provides the empirical results and estimates the premiums for different majors. Section 5 examines how well the model predicts the trends seen in the data. Section 6 simulates how the ability sorting across majors would change given changes in the environment. Section 7 concludes.

2. Education choices and earnings outcomes

This section provides descriptive statistics on the earnings and characteristics of individuals who participated in the National Longitudinal Study of the Class of 1972 (NLS72).[2] Only those who indicated that they had been accepted to a 4-year college were included in the sample as I am only interested in those for whom attending a

---

Table 1
1986 Earnings[a] by college major

Major	1972 choice	1974 choice	Difference
Natural science	50,535	52,315	1780
	(24,805)	(24,419)	
Business	49,249	52,796	3547
	(26,227)	(23,015)	
Social science/	38,955	43,088	4133
humanities	(18,583)	(22,288)	
Education	33,616	32,305	−1311
	(13,589)	(10,417)	
No college	36,478	36,664	185
	(18,016)	(18,642)	

[a]Translated into 1999 dollars. Standard deviations in parentheses.

4-year college was a serious consideration.[3] I also aggregate majors into four categories: Natural Sciences (including math and engineering), Business (including economics), Social Science/Humanities/Other, and Education. The criteria for aggregation was the degree of similarity in mean earnings and SAT math and verbal scores.[4]

With these assumptions, the major findings of the descriptive analysis are:

1. Earnings are strongly correlated with major choice.
2. Ability sorting across majors occurs both before and during college.
3. Lucrative majors draw the high math ability students at each school.
4. Poor performance is correlated with dropping out or switching to a less lucrative major.

Table 1 addresses the first point listed above, displaying 1986 mean earnings data for individuals by their intended major going into college (1972 college

---

[3] A significant fraction of individuals who are accepted to a 4-year college decide not to attend. All individuals had to have test score information to stay in the sample. Test information could be either the SAT or the test given to participants in the first round of the survey. For the latter, I use the NLS72 tests to predict their SAT scores. To stay in the sample and choose a schooling option required that grades be observed for the next year, that the individual attend one of his top three schooling choices, and that, if the individual did not choose a schooling option in 1972, he also did not choose a schooling option in 1974. These latter assumptions are made to satisfy the confines of the model.

[4] The aggregation has little effect on the descriptive trends presented here. For example, Social Science and Humanities majors look very similar in both math ability and earnings as do physical science majors and engineers. The one piece of evidence that is lost here is that biology majors are more likely to have lower math abilities and more likely to be women than other majors in the natural science category. They are aggregated here to keep the cell counts high.

major) and their major halfway through their college career (1974 college major). [5]

Note the more than $16,000 dollar mean spread between the highest paying major, natural science, and the lowest paying major, education, for the 1972 choice. In fact, those who chose not to attend college actually had higher average earnings than those who chose education either in 1972 or in 1974.

The difference in earnings between the natural sciences and education increases once the students update their major decision. With the exception of education, all majors see higher earnings after the re-sorting. [6] The re-sorting implies that, with the exception of education, real premiums exist across majors and/or high ability individuals are more likely to choose the more lucrative fields later in college. Table 2 provides some evidence for the latter.

Table 2 shows mean math and verbal SAT scores by major as well as the corresponding means for the individual's peers for both 1972 and 1974. [7] The abilities of the individual's peers are defined as the mean ability level at the college or university; [8] no measure of peer ability is available at the major level. With the differences in math abilities across majors, we begin to see that the results in Table 1 may be in large part due to selection. [9] Differencing one's own ability from the ability of his peers shows that the ordering of majors by earnings is very similar to the ordering by relative math ability. This is not the case when we add the verbal score as business majors do substantially worse than their peers on the verbal portion of the SAT. In fact, while business majors have essentially the same math abilities and attend the same quality institutions as the social science and humanities majors, their verbal scores are much lower. It is interesting to note that natural science majors not only have the highest math SAT scores, but the highest verbal scores as well.

Selection also plays a role in who stays in college. Table 2 shows that for all majors both average student abilities and average peer abilities increase after allowing students to drop out of college or switch to a different major. The largest differences across years in average abilities were in the most lucrative majors. Dropouts had lower math and verbal scores and attended colleges with low average abilities. This can explain part of the increases in average earnings for non-education majors between the 1972 and 1974 choice. Education majors in 1972 look very similar to education majors in 1974,

---

[5] Individuals must have worked between 30 and 60 h in an average week and, given that average work week, earned between five and a half and one hundred and forty eight and a half thousand dollars a year (1999 dollars). Including those who work less than 30 h or who are unemployed would exacerbate the trends as those who choose natural science majors are more likely to be employed.

[6] Although 1974 business majors are listed here as having higher earnings than natural science majors, under less rigorous selection rules this does not hold. This is the only case where the qualitative results change because of the selection rules.

[7] For those who did not take the SAT, SAT scores were predicted using scores from a standardized test taken by the survey recipients and demographic characteristics.

[8] This data is taken from the institutions themselves.

[9] Although we would expect the average SAT math and verbal gaps to be zero, the actual average math and verbal gaps are −15.1 and −25.7, respectively. If we limit the sample to those who take the SAT, the problem is somewhat mitigated shrinking the math and verbal gaps to −6.3 and −18.3, respectively.

Table 2
SAT scores by college major

	Major	1972 choice			1974 choice		
		Own ability	Peer ability	Relative ability	Own ability[a]	Peer ability	Relative ability
SAT math	Natural science	566 (103)	547 (62)	19 (93)	594 (98)	560 (64)	34 (88)
	Business	498 (105)	522 (58)	−24 (90)	528 (92)	533 (55)	−5 (87)
	Social science/ humanities	500 (104)	526 (58)	−26 (95)	518 (100)	535 (56)	−17 (91)
	Education	458 (95)	502 (50)	−44 (89)	467 (95)	504 (54)	−37 (84)
	No college[b]	430 (102)			482 (110)	514 (57)	−32 (102)
SAT verbal	Natural science	499 (106)	515 (58)	−16 (96)	523 (106)	526 (60)	−3 (93)
	Business	444 (96)	494 (57)	−50 (85)	464 (93)	501 (52)	−38 (90)
	Social science/ humanities	481 (107)	499 (56)	−18 (96)	499 (102)	510 (56)	−12 (93)
	Education	431 (92)	477 (48)	−46 (87)	438 (89)	478 (51)	−40 (84)
	No college	404 (98)			445 (106)	486 (53)	−41 (97)

[a] For those who did not choose a college option in 1974, peer ability is based upon their 1972 choice.
[b] No college in 1974 refers to those who attended in 1972 but not in 1974. 'Relative Ability' is own ability minus peer ability. Standard deviations in parentheses.

suggesting that whatever increases in earnings we would expect to see by removing dropouts is offset by individuals switching into and out of education.

Table 3 provides some information on who it is that is transferring into and out of particular majors or dropping out. Significant differences exist in drop-out rates across majors, with natural science and business majors dropping out less often than education and social science/humanities majors. Despite the low drop-out rates, natural science majors are the least likely to stay with their initial choice. This suggests that it may be easier to switch into other majors from a natural science major.

The characteristics of those who drop out are very different across initial majors. While business majors drop out at the same rate as natural science majors, the abilities

Table 3
SAT and earnings transitions

1972 major	Variable	1974 major					
		Natural science	Business	Social science/ hum	Education	Drop out	Total
Natural science	Own math ability	602	555	543	494	537	566
	Peer math ability	565	539	539	495	532	547
	Own verbal ability	529	468	496	461	466	499
	Peer verbal ability	531	504	514	472	498	515
	1986 earnings	54,333	55,348	47,309	33,849	44,473	50,535
	% of '72 Natural science	42%	8%	19%	2%	28%	100%
Business	Own math ability	579	516	506	498	455	498
	Peer math ability	534	530	539	515	502	522
	Own verbal ability	449	452	470	439	420	444
	Peer verbal ability	495	501	515	489	475	494
	1986 earnings	42,860	51,862	55,700	38,110	40,429	49,249
	% of '72 Business	3%	54%	10%	3%	30%	100%
Soc sci/ humanities	Own math ability	546	533	515	475	474	500
	Peer math ability	530	541	535	515	513	526
	Own verbal ability	501	493	503	456	452	481
	Peer verbal ability	504	508	509	484	487	499
	1986 earnings	40,273	51,211	41,855	31,289	33,621	38,955
	% of '72 Soc sci/hum	4%	4%	49%	7%	36%	100%
Education	Own math ability	583	537	479	460	433	458
	Peer math ability	544	525	513	500	495	502
	Own verbal ability	495	492	478	429	409	431
	Peer verbal ability	516	495	490	476	471	477
	1986 earnings	51,176	57,437	36,537	32,315	30,364	33,616
	% of '72 Education	2%	3%	11%	51%	33%	100%

of the business major dropouts and their peers are substantially below their natural science counterparts. Despite this, earnings for business major dropouts are only slightly lower than earnings for natural science dropouts and significantly higher than social science/humanities majors. Besides having much lower verbal scores, business major dropouts look very similar to social science/humanities majors. Those who drop out of education are of significantly lower abilities, attend lower quality schools, and earn much less than all of the other drop-out groups. Further, conditional on choosing a

Table 4
Grades in 1972 major by switching category[a]

	Natural science	Business	Social science/ humanities	Education	Total
Switch up		2.71	2.84	2.94	2.87
		(0.39)	(0.60)	(0.56)	(0.59)
Same major	3.03	2.71	3.00	2.78	2.93
	(0.64)	(0.58)	(0.56)	(0.54)	(0.59)
Switch down	2.69	2.59	2.82		2.71
	(0.64)	(0.70)	(0.56)		(0.63)
Drop out	2.51	2.44	2.60	2.51	2.55
	(0.70)	(0.76)	(0.65)	(0.70)	(0.69)

[a]Switching down refers to switching to a less lucrative (low SAT math gap) major, while switching up refers to switching to a more lucrative (high SAT math gap) major.

major besides education, dropouts earn more than those who switch into education. This may be a result of the abilities of the individuals who switch into education or through non-pecuniary benefits associated with being an education major. In all other cases, dropouts earn less.

Comparing cross major switches (those who choose major i then major j with those who choose j then i) yields more ability sorting. Natural science majors who switch to business have on average lower math scores than business majors who switch to the natural sciences. Similarly, those who switch from natural science to social science/humanities or education have lower math abilities than the social science/humanities and education majors who switch to the natural sciences. Excluding natural science majors, business majors trade lower math ability students for higher math ability students as well. Finally, social science/humanities majors who switch to education are of lower math abilities than education majors who switch to social science/humanities. [10] There is also some evidence of comparative advantage as those who have high math scores relative to verbal scores are more likely to choose a natural science major over a social science/humanities major.

With over 30% of those attending college in 1972 dropping out by 1974 and another 18% switching majors, [11] individuals must be learning about their tastes and abilities. Table 4 presents some evidence that learning about one's abilities affects one's future educational decisions. Table 4 displays freshmen grade point averages [12] and standard deviations by 1972 major across four groups: those who switched to a more lucrative

---

[10] The same pattern is observed for verbal scores with one exception: business majors who switch to the natural sciences score worse than natural science majors who become business majors.

[11] By definition, disaggregating the major categories further would lead to even more switching.

[12] This is a categorical variable which is taken from the survey. Students were asked to give their average g.p.a. at the time of the survey. Midpoints of the categories were used in all mean calculations. I use freshmen grade point averages to get grades closest to the initial decision.

major (switched up), stayed in the same major, switched to a less lucrative major (switched down), and those who dropped out.

In all cases those who dropped out had lower grades than the other three categories. Also, those who switched down performed worse than those who stayed in the same major or switched up. [13] Comparing switching up to keeping the same major leads to a more muddled picture. In two of the three cases those who switched up had performed better than those who stayed the same, though this pattern is not observed in the largest category. It may be that how lucrative the major is does not perfectly match with difficulty—switching from a social science major to a business major may not be a switch to a more difficult major. This suggests that there may be a major-specific component to what is learned in college. Those who perform poorly know to try a different major or drop out. If performing well is more important in the lucrative majors, poor performers who stay in school will find one of the lower paying majors more attractive. Those who perform well have an incentive to stay in the same major to take advantage of the major-specific skills. However, to the degree that the discovered abilities are general, students have an incentive to switch to majors where their abilities may be put to better use.

3. Model and estimation

The trends in the data suggest a dynamic model of college decision making. I model the college education process as consisting of three periods. In the first two periods individuals decide among a variety of schooling options or choose to enter the workforce. All individuals work in the third period, reaping the benefits of their past educational decisions. A broad outline of the model is summarized below.

1. In period 1, individuals are given a choice set from which they can choose both a college and a major or enter the labor market. The choice set is the set of schools where the individual was accepted. The labor market is an absorbing state.
2. After the first period decision, those who chose a schooling option receive new information about their abilities (through grade point averages) and how well they like particular fields (through preference shocks). [14]
3. In period 2, those who pursued a schooling option again choose from the same schooling options as in period 1 or enter the labor market.
4. After the second period decision, individuals who chose a schooling option in the second period again receive new information about their abilities. They then enter the labor market in period 3.

Periods 1 and 2 roughly correspond to the individual's first two and last two years of college. Period 3 includes all years after college. The model involves estimating

---

[13] Grades for the dropout category as well as grades for the switch down category are significantly different from each other and significantly different from the other two categories at the 95% level.

[14] Learning about preferences may also affect grades if individuals discover they do not like a particular field mid-year and then choose not to work as hard in their classes.

parameters of two types: utility function parameters (α's) and transition parameters (γ's). Transition parameters are used only in forming expectations about uncertain future events. These include the parameters of the grade generating process, through which individuals learn about their abilities and the corresponding value of pursuing particular educational paths, and the parameters of the earnings process which dictate expectations individuals have about future financial outcomes. I first discuss the transition parameters and for the moment assume that errors are independent across all stages of the model. I then relax this assumption and allow the errors to be correlated through the use of mixture distributions later in the paper. Although the model is estimated using individual-level data, the individual i subscripts are omitted throughout to simplify the notation.

3.1. The labor market

Once individuals enter the workforce they make no other decisions: the labor market is an absorbing state. Earnings are a function of observed ability, A, where A is individual specific. I assume that the human capital gains for attending the jth college operate through the average ability of the students at the college, $\bar{A}_j$, grades received, G, and the major chosen. I assume that log earnings in year t are given by

$$\log(W_t) = \gamma_{w1k}A + \gamma_{w2k}\bar{A}_j + \gamma_{w3k}G_{jk} + \gamma_{w4k}Z_w + g_{wkt} + \varepsilon_{wt}, \tag{1}$$

where individual i subscripts are suppressed. The subscript k refers to the major chosen. G_{jk} then represents cumulative grade point averages in college conditional on choosing school j and major k. Z_w is a vector of other demographic characteristics, such as gender, which may affect earnings. g_{wkt} is the major-specific time trend on earnings. The shocks (the ε_{wt}'s) are assumed to be distributed $N(0, \sigma_w^2)$. The subscript w refers to coefficients or variables which are a part of the data generating process for earnings.

I use the SAT math and verbal scores as my measures of observable ability. The corresponding school averages [15] are used for the measure of school quality. Average abilities need not be interpreted as peer effects but as the best measure of college quality available. Crucial to identification of the coefficient on earnings in the utility function is that an exclusion restriction exists; a variable which appears in the earnings equation and only affects utility through earnings. I use average state earnings for both workers who graduated from college and those who did not graduate from college as the exclusion restriction. This variable is calculated from the March Current Population Survey (CPS). [16]

---

[15] These data are taken from the colleges themselves. The Basic Institutional Source File has information taken from the 1973–1974 Higher Education Directory, the 1973–1974 TriparIate Application Data file, the 1972–1973 HEGIS Finance Survey and the 1972 ACE Institutional Characteristics File.

[16] Sparsely populated states are aggregated in the CPS, so instead of 50 data points there are actually only 22. I regress log earnings for those who are 22–35 years of age on an age quadratic for both men and women. I then pull out the gender and age specific effects and average across regions to obtain the college premiums. I use the same restrictions on extreme observations as in the previous section.

3.2. Learning

While grade point averages are expected to have a positive effect on future earnings, individuals learn about their abilities through them as well. A signal on unobserved ability is given in the realization of first period college performance. This unobserved ability, A_{uk}, may be partially major specific. Specifically, let A_{uk} follow:

$$A_{uk} = \eta_1 + \eta_{2k}.$$

I assume η_1 is distributed $N(0, \sigma_{u1}^2)$ while η_{2k} is distributed $N(0, \sigma_{u2}^2)$ for all k. The first term (η_1) is the portion of learned ability which is transferable across majors and is therefore independent from the second term (η_{2k}), the portion of learned ability which is major-specific. Since this is ability which individuals were not able to forecast, it is independent from all information individuals have before the realization on grades.

In addition to a noisy signal on unobserved ability, performance in the first period, G_1, is a function of the major chosen, the individual's own abilities and the abilities of his classmates, as well as Z_g which represent other demographic characteristics such as high school class rank and gender. Specifically, performance in the first period takes the following form:

$$G_1 = \gamma_{g1k} + \gamma_{g2k}A + \gamma_{g3k}\bar{A}_j + \gamma_{g4k}Z_g + A_{uk} + \varepsilon_{g1}, \tag{2}$$

where γ_{g1} is a vector of coefficients to be estimated and ε_{g1} is a white noise component distributed $N(0, \sigma_{g1}^2)$.

Each individual takes the difference between his actual performance and expected performance as a signal on his unobserved ability. That is, I assume the econometrician and the individual have the same information set when predicting first period performance, an assumption which will be relaxed when controls for unobserved heterogeneity are implemented later in the paper. Let the signal given on the unobserved ability be called A_{u1k}, where $A_{u1k} = A_{uk} + \varepsilon_{g1}$.

The unobserved ability also affects second period performance, which in turn affects the present value of lifetime earnings in the final period. I assume the second period performance process has the same parameters as the first period performance process, up to a parameter on the intercept term and on the ability terms. That is, while the mean grades across majors may change over time, the relative importance of math and verbal ability does not change from one period to the next. Performance in the second period then takes the following form:

$$G_2 = \gamma_{g5}\gamma_{g1k} + \gamma_{g6}(\gamma_{g2k}A + \gamma_{g3k}\bar{A}_j + \gamma_{g4k}Z_g) + \gamma_{g7}S + \gamma_{g6}A_{uk} + \varepsilon_{g2}, \tag{3}$$

where S indicates that the individual switched majors and ε_{g2} is again a white noise component and is distributed $N(0, \sigma_{g2}^2)$. Individuals then use the information they receive from first period performance to forecast second period performance according to:

$$E_1(G_2|j,k) = \gamma_{g5}\gamma_{g1k} + \gamma_{g6}(\gamma_{g2k}A + \gamma_{g3k}\bar{A}_j + \gamma_{g4k}Z_g) + \gamma_{g7}S$$

$$+ \gamma_{g6}\frac{\sigma_{u1}^2 + \sigma_{u2}^2 S}{\sigma_{u1}^2 + \sigma_{u2}^2 + \sigma_{g1}^2} A_{u1k}. \tag{4}$$

Define the forecast error as ε_g. ε_g is then the sum of normally distributed variables which are independent from Z_g.[17] The forecast errors are heteroscedastic as those who switch majors are expected to have higher variances on their error terms. The signal to noise ratios for those who stay in the same major versus those who switch can then be used to identify the importance of major specific ability versus absolute ability. In particular, if the signal to noise ratios are the same for stayers and switchers then there is no learned ability which is major-specific. However, large differences in the signal to noise ratios would indicate a strong presence of major-specific ability.

For my performance measure, I use the individual's college grades during the year immediately after the student has made the period 1 and period 2 choices. By focusing on grades immediately after the decisions, I hope to mitigate the effect on grades of those who switch majors in between the two periods. Hence, the grades used will be those reported in 1973 and 1975. All school variables are based upon the choices made in October of 1972 and 1974.

3.3. Choice of college and major

Individuals may choose a school from a set J where colleges themselves are not important; it is only the characteristics of the colleges that are relevant to the model. That is, one does not receive utility from attending Harvard but from attending a school that has faculty and students with particular characteristics. Any effect that Harvard itself has must be captured by these characteristics. Those who decide to attend college must also choose a major from the set K. I assume that the same set of majors exists at all colleges.[18] When making the college and major decisions, individuals take into account the repercussions these decisions have on future earnings.

The NLS72 has data on the top three schooling choices of the individual in 1972 and on whether or not the individual was accepted to each of these schools. It also has data on the schooling choice made in 1972 and 1974 and I restrict the data set to those students who attend one of their top three choices in both periods.[19] Unfortunately, the NLS72 does not have data on whether an individual was considering any other 4-year institutions. Hence, I may only be partially observing the choice set.[20] I aggregate majors into four categories as in the previous section. The maximum number of choices available in periods 1 and 2 is then 13: four majors for each of three schools and a work option.

---

[17] To see this, consider the regression $A_{uk} = \gamma A_{u1k} + \varepsilon$. The error from this regression is, by construction, orthogonal to A_{u1k}. Z_g is then also orthogonal to ε since the only correlation Z_g had with A_{uk} was through the sum A_{u1k}. γ is then the signal to noise ratio: how much information the draw on A_{u1k} is providing on A_{uk}.

[18] I avoid the issue of small liberal arts schools having much fewer majors through the aggregation of majors into broad categories. The assumption is then that each school offers at least one major in each of the broad categories.

[19] Individuals transferring to school outside of the original three is not allowed and these individuals are removed from the data set. Clearly, one's performance in college may dictate what schools an individual can transfer to. Unfortunately, data does not exist as to what transfer options were available to the individual.

[20] Not observing other schools in the choice set does not appear to be important as those students who applied to at least three schools are less than 15% of all NLS72 participants who applied to college.

I assume that utility is separable over time. Utility of being in the workforce is a function of the log of the expected present value of the of lifetime earnings as well as preferences individuals have for the jobs associated with particular majors net of the monetary returns. Utility in the workforce then follows:

$$u_{wjk} = \alpha_{w1k}A + \alpha_{w2k}\bar{A}_j + \alpha_{w3k}X_{wj} + \alpha_{w4} \log \left(E_w \left[\sum_{t=t'}^{T} \beta^{t-t'} P_{kt} W_t \right] | j, k \right), \qquad (5)$$

where T is the retirement date, t' is the year the individual enters the workforce, and P_{kt} is the probability of working in year t given major k. The first three terms represent the preference components to working in a particular field. For example, an individual who is not good at math may not want a math-intensive job beyond the fact that he may be compensated less because of his poor math skills. The expression for utility can then be rewritten as: [21]

$$u_{wjk} = (\alpha_{w1k} + \alpha_{w4}\gamma_{w1k})A + (\alpha_{w2k} + \alpha_{w4}\gamma_{w2k})\bar{A}_j + \alpha_{w3k}X_{wj}$$
$$+ \alpha_{w4}(\gamma_{w3k}G_{jk} + \gamma_{w4k}Z_w) + \alpha_{w4}E \left(\log \left[\sum_{t=t'}^{T} \beta^{t-t'} P_{kt} \exp(g_{wkt} + \varepsilon_{wt}) \right] \right).$$
$$(6)$$

I assume that all individuals have the same expectations on the probabilities of working in particular years conditional on sex and major. This, combined with the assumptions made on the growth rates and earnings shocks, imply that the last term is not separately identified from the major-specific intercepts and the corresponding interactions with gender. This is a fortunate product of using the log specification: forecasting the probability of working and earnings growth rates far out into the cycle (and well beyond where the data lie) is then not needed. Crucial, then, to identification of the coefficient on earnings (α_{4w}) in the utility function is that an exclusion restriction exists; a variable in Z_w which is not in X_w. As discussed in Section 3.1, I use average state earnings for both workers who graduated from college and those who did not graduate from college as the exclusion restriction.

Define the flow utility u_{tjk} as the utility received while actually attending college j in major k at time t. This flow utility includes the effort demanded for the particular schooling combination, c_{jk}, as well as any preferences individuals have for particular majors at particular schools. The flow utility for pursuing a particular college option is then:

$$u_{tjk} = \alpha_{1k}A + \alpha_{2k}\bar{A}_j + \alpha_{3k}X_j - c_{jk} + \varepsilon_{tjk}, \qquad (7)$$

---

[21] One of the advantages of choosing the log utility specification is that errors in growth rates result in changes in the coefficients on the constant and gender terms in the utility specification but do not affect other parameter estimates. While the NLS72 has good data on yearly earnings for 1973–1979 and also for 1986, we have little information on the growth rates by major late into the life cycle.

where X_j is a vector of individual and school characteristics which affect how attractive particular education paths are.[22] The individual's unobserved preferences for the schooling options is given by the ε_{tjk}'s.

Each of the majors varies in its demands upon the students. I assume that each major requires a fixed amount of work which varies by the individual's ability, A, ability of one's peers, $\bar{A}_j$, the ability that is learned about in college, A_{uk}, and the major chosen, k. Let A_M and A_E represent math and verbal ability, with total ability given by $A_T = A_M + A_E$. In order to conserve on parameters, the cost of effort from the observed abilities, c_{jk}, is assumed to follow:

$$c_{jk} = \alpha_{c1k}(A_M - \bar{A}_{jM}) + \alpha_{c2k}(A_V - \bar{A}_{jV}) + \alpha_{c3}(A_T - \bar{A}_{jT})^2. \tag{8}$$

Note that the psychic cost function allows the costs to majoring in particular fields to vary by relative ability in the linear term, but not in the squared term. This cost of effort may lead to optimal qualities that are on the interior; even if an individual was allowed to attend all colleges, the individual may choose not to attend the highest quality college because of the effort required. With different levels of effort required by different majors, optimal college qualities may vary by major. At the margin, individuals are then trading off the cost of obtaining the human capital with the future benefits.

While the coefficient on the squared term is identified, the other parameters of the cost of effort function are not. The reason is that preferences which vary across abilities may exist for particular fields irrespective of the cost of effort and the quality of the college may serve as a consumption good. Additional assumptions, outlined in Section 3.4, are necessary to identify the cost function.

After making the second period college decision, there are no decisions left and the individual enters the workforce. The expected present discounted value of lifetime utility conditional of choosing a college option in the second period, v_{2jk}, is then given by

$$v_{2jk} = E_2(u_{2jk} + \beta u_{wjk}), \tag{9}$$

where β is the discount rate. Individuals then choose the option which yields the highest present value of lifetime utility. Note that the unobservable preference term ε_{2jk} is embedded in u_{2jk} and is known to the individual but not the econometrician. What is unknown to the individual are grades in the second period and the time path of earnings. The expected present value of choosing to enter the workforce (as opposed to choosing one of the schooling options) is then the sum of the expected log of lifetime earnings, u_{wo}.[23]

In the first period, individuals take into account how their actions will affect the value of their future choice. Let V_2 indicate the best option in the second period. Individuals

---

[22] I have no data on major specific variables, hence there is no k subscript. In addition to allowing preferences for particular majors to vary based upon individual and peer ability as well as gender, the cost of the school (tuition plus room and board) net of scholarships, whether the school is private and whether the school is in the same state are also allowed to influence whether and individual attends college.

[23] Any preferences in the workplace that vary across abilities for particular educational outcomes are then relative to the preferences that vary across abilities for working without a college degree.

then choose the v_{1jk} which yields the highest utility where v_{1jk} is given by

$$v_{1jk} = E_1(u_{1jk}) + \beta E_1(V_2|d_1 = j, k). \tag{10}$$

This second expectation is taken with respect to both shocks to ability and shocks to preferences. Individuals get to optimize again after the realization of these shocks, but there is a cost to not knowing this information a priori.

By integrating out the new information on one's abilities in the expectation of future utility and assuming that the new information is uncorrelated with the unobservable preferences, Eq. (11) results:

$$v_{1jk} = E_1(u_{1jk}) + \beta \int E_1(V_2|A_{u1}, d_1 = j, k)\pi_1(\mathrm{d}A_{u1k}|X_j, d_1 = j, k), \tag{11}$$

where π_1 is the pdf of the signals on unobserved ability. Note that there is still an expectation operator in front of the future utility component because individuals receive draws on their unobservable preferences after making their first period decisions. Even if the new information on ability was known to the individual, the second period decision would still be stochastic because of the evolution of the unobservable preference parameters (the ε's).

In order to actually estimate models of this type, some assumptions need to be made on the distribution of the unobserved preferences. Specifically, let the ε_{tjk}'s in each time period be taken from a generalized extreme value distribution which yields nested logit probabilities in a static model: schooling options in one nest, work options in the other. That is, ε_{tjk} has a component which does not vary across schooling options. Let the variance for the cross-school component at time t be given by μ_{2t}. The variance on ε_{tjk} itself is given by μ_{1t} where μ_{1t} must be greater than μ_{2t}.[24]

With the added assumption that the unobservable preference terms are uncorrelated over time,[25] closed form solutions for the conditional expectations of future utility exist. Specifically, the present value of indirect utility for attending school j in major k during period 1 is now given by Eq. (12):[26]

$$v_{1jk} = E_1(u_{1jk}) + \beta\mu_{12}\int \ln\left[\left(\sum_j \sum_k \exp[(E_1 v_{2jk}|A_{u1k}, d_1 = j, k)]\right)^{\mu_{22}/\mu_{12}}\right.$$

$$\left. + \exp(u_{wo})\right]\pi_1(\mathrm{d}A_{u1k}). \tag{12}$$

Recall that A_{u1k} is found through the first period performance regression given in Eq. (2), where A_{u1k} was assumed to be a normally distributed random variable. In order to evaluate this expression Rust (1987) showed that, by discretizing the values A_{u1k}

---

[24] All parameters in discrete choice models are relative to the variance term where the variance term is usually normalized. Here, I normalize with respect to μ_{11}.

[25] An assumption which is made more palatable in the next section.

[26] See McFadden (1981) for the result.

can take, it is possible to obtain consistent estimates of the parameters of the utility function.[27] With $p_1(\cdot)$ being the discretized version of $\pi_1(\cdot)$, Eq. (13) results.[28]

$$v_{1jk} = E_1(u_{1jk}) + \beta\mu_{12} \sum_m \ln \left[\left(\sum_j \sum_k \exp[E_1(v_{2jk}|A_{u1km}, d_1 = j, k)] \right)^{\mu_{22}/\mu_{12}} \right.$$

$$\left. + \exp(u_{wo}) \right] p_1(A_{u1km}). \tag{13}$$

With the assumptions made on the distribution of the unobservable preferences and the earnings and grade processes, the probability of an individual choosing school j and major k in period t takes a nested logit form:[29]

$$Pr(d_t = j, k) = \frac{\exp(v'_{tjk})}{\sum_j \sum_k \exp(v'_{tjk}) + (\sum_j \sum_k \exp(v'_{tjk}))^{1-\mu_{2t}} \exp(u_{wo})},$$

where the sums are taken over all possible options available to the individual. The expected net present value of indirect utility for attending school j in major k during period t is then given by v'_{tjk}. It is 'net' because I am differencing the present value of lifetime utility for educational choice in period t by the present value of pursuing the work option in period t. It is 'expected' for two reasons. First, both the researcher and the individual only have expectations regarding the value of future decisions. It is also 'expected' because the unobservable preferences are unobserved to the researcher, and I am defining v'_{tjk} such that it does not include these unobservable preferences. Since only part of the indirect utility is observed in v'_{tjk}, the decision an individual makes from the researcher's standpoint is random.

3.4. Restrictions on preferences in college and the workplace

Given three ability variables (SAT math, SAT verbal, and high school class rank), two college quality variables (average SAT math and average SAT verbal), preferences in the workplace and at college, and four majors, forty preference parameters exist just from ability and college quality. All of these measures are highly correlated. Assumptions need to be made to restrict the number of parameters estimated. Further, without assumptions on the preferences it will be impossible to separate out preferences for college quality from effort costs associated with attending higher quality schools. The assumptions I make nest both models that do not have effort costs and that do not

---

[27] Keane and Wolpin (1994) present an alternative method which does not involve discretizing the error term. Their method involves approximating the integrals with different functions. Keane and Wolpin (1997) use this method in their model of career decisions. For good reviews of solution methods for dynamic discrete choice problems see Eckstein and Wolpin (1989) and Rust (1994, 1996).

[28] Here, A_{u1k} does not depend upon X_j as the expectations are on the forecast error which are independent from X_j in the first period.

[29] Rust (1987) showed this explicitly for the multinomial logit case and his general model produces the nested logit specification as a specific case.

have preferences for particular majors which vary with ability or college quality either in college or in the workplace.

The first assumption restricts the preference variation across individuals with different abilities. I assume that the degree to which high ability individuals have preferences for particular majors is proportional to the job satisfaction high ability individuals have in the careers associated with those majors. That is, I restrict the utility function parameters from Eqs. (5) and (7) for each of the K majors such that:

$$\alpha_{w1k} = \phi_1 \alpha_{1k}. \tag{14}$$

The second assumption restricts the pattern of preference variation for different college qualities both in an out of the workplace. In particular, I assume that the degree to which high-quality colleges make particular majors more attractive while in college and in the workplace is proportional to the preferences high-ability individuals have for particular fields. That is, I restrict the preference parameter from Eqs. (5) and (7) such that both:

$$\alpha_{2k} = \phi_2 \alpha_{1k}, \tag{15}$$

$$\alpha_{w2k} = \phi_1 \phi_2 \alpha_{1k}. \tag{16}$$

These assumptions reduce the number of parameters from 40 to 42. The effort costs, which previously could not be separated from the other preference parameters, are now identified. Identification comes from patterns of behavior which are not proportional. No restriction is placed on the sign of the effort costs; the model may produce estimates which are inconsistent with the theory. I test these restrictions by estimating the unrestricted and restricted models and using a likelihood ratio test.

These assumptions also admit models with no preference variation across abilities and college quality, either in college or in the workplace, as special cases. For example, if ϕ_1 equals zero, then there is no preference variation across abilities in the workplace; the choice of major is not tied to job characteristics. Further, if ϕ_2 equals zero, then college quality is not a consumption good.

3.5. Estimation strategy

With independent errors across the earnings, grades, and choice processes, the log likelihood function is the sum of three pieces:

$L_1(\gamma_w)$—the log likelihood contribution of earnings,
$L_2(\gamma_g)$—the log likelihood contribution of grade point averages,
$L_3(\alpha, \gamma_g, \gamma_w)$—the log likelihood contribution of college and major decisions,
with 143 parameters to estimate.

It is possible to estimate all the parameters in the indirect utility function, the performance equations, and the log earnings equations using full information maximum likelihood. However, this would be computationally burdensome. Note that consistent estimates of γ_w and γ_g can be found from maximizing L_1 and L_2 separately.[30] With

---

[30] See Rust and Phelan (1997) and Rothwell and Rust (1997).

the estimates of γ_w and γ_g, consistent estimates of α can be obtained from maximizing L_3. [31]

3.6. Serial correlation of preferences and unobserved ability

One of the assumptions which seems particularly unreasonable is that the unobservable preferences parameters are uncorrelated over time. That is, a strong unobservable preference for engineering is not associated with having a strong unobservable preference for engineering in the second period. We would suspect that this is not the case. Further, it is unreasonable to assume that there is no unobserved (to the econometrician) ability which is known to the individual. [32]

Mixture distributions provide a way of controlling for serial correlation and selection. Assume that there are R types of people with π_r being the proportion of the rth type in the population. [33] Types remain the same throughout all stages, individuals know their type, and preferences for particular fields and college quality may then vary across types. An example would be if the parameters of the utility function do not vary across types except for the constant term. This would be the same as having a random effect which is common across everyone of a particular type. The log likelihood function for a data set with I observations is then given by

$$L(\alpha, \gamma_g, \gamma_w) = \sum_{i=1}^{I} \ln \left(\sum_{r=1}^{R} \pi_r \mathscr{L}_{1ir} \mathscr{L}_{2ir} \mathscr{L}_{3ir} \right). \tag{17}$$

Here, the α's and γ's can vary by type and $\mathscr{L}$ refers to the likelihood (as opposed to the log likelihood).

Now the parts of the log likelihood function are no longer additively separable. If they were, a similar technique could be used as in the case of complete information: estimate the model in stages with the parameters of previous stages being taken as given when estimating the parameters of subsequent stages. Using the EM algorithm, [34] I am able to return the additive separability.

Note that the conditional probability of being a particular type is given by

$$Pr_i(r|\mathbf{X_i}, \alpha, \gamma, \pi) = \frac{\pi_r \mathscr{L}_{1ir} \mathscr{L}_{2ir} \mathscr{L}_{3ir}}{\sum_{r=1}^{R} \pi_r \mathscr{L}_{1ir} \mathscr{L}_{2ir} \mathscr{L}_{3ir}}, \tag{18}$$

where $\mathbf{X_i}$ refers to the data on the decisions and the characteristics of the individual.

---

[31] The standard errors are not consistent, however, unless the covariance matrix of the parameters is *block diagonal* as estimates of transition parameters are being taken as the truth. Full information maximum likelihood with one Newton step would produce consistent estimates of the standard errors. See Davidson and MacKinnon (1993) for how using this two-step method affects the standard errors. Rust and Phelan (1997) note that in other work the two stage estimation procedure has had little effect on the standard errors.

[32] See Willis and Rosen (1979) for the importance of selection in education.

[33] See Keane and Wolpin (1997) and Eckstein and Wolpin (1999) for other examples of using mixture distributions to control for unobserved heterogeneity in dynamic discrete choice models.

[34] See Dempster et al. (1977).

The EM algorithm has two steps: first calculate the expected log likelihood function given the conditional probabilities at the current parameter estimates, second maximize the expected likelihood function holding the conditional probabilities fixed. This process is repeated until convergence is obtained. But the expected log likelihood function here is now additively separable.

$$\sum_{i=1}^{I} \sum_{r=1}^{R} Pr_i(r|\mathbf{X_i}, \alpha, \gamma, \pi)(L_{1ir}(\gamma_w) + L_{2ir}(\gamma_g) + L_{3ir}(\alpha, \gamma_g, \gamma_w)). \tag{19}$$

Taking the conditional probabilities as given, I can get estimates of γ_w from maximizing the L_{1r}'s times the conditional probabilities. Similarly, estimate γ_g from maximizing the conditional probabilities times the L_{2r}'s. I then only use the L_{3r}'s to find estimates of α—not needing the L_{3r}'s to obtain estimates of γ_w and γ_g. Note that all of the parts of the likelihood are still linked through the conditional probabilities where the conditional probabilities are updated at each iteration of the EM algorithm. Arcidiacono and Jones (2003) show this method produces consistent estimates of the parameters with large computational savings.

4. Empirical results

This section provides the results from estimating the parameters of the performance equations, the log annual earnings equations, and the structural parameters of the utility function. Although the results of the model with unobserved heterogeneity are interdependent, I present the estimation of each equation separately.

4.1. Performance regressions

The results of estimating the parameters of the first period performance equation are given in Table 5. The first column displays the coefficient estimates without unobserved heterogeneity, while the second gives estimates with unobserved heterogeneity approximated by two types as discussed in Section 3.6. Two additional restrictions are placed on the coefficients. First, the coefficient on math (verbal) college quality is constrained to be proportional to the coefficient on math (verbal) ability. The sign, however, is not constrained. Second, for some individuals we do not know their class rank. The coefficient on 'do not know class rank' is constrained to be proportional to the effect of high school class rank. The assumption is then that the coefficient on 'do not know class rank' yields what we would expect their class rank to be.

All of the ability coefficients are positive, with smaller coefficients for education. Without unobserved heterogeneity, math ability is particularly useful in the natural sciences, while verbal ability is particularly useful in the social science/humanities. Once the mixture distribution is added, the differences in ability coefficients within a major dissipate. High school class rank positively affects grade point averages. Calculating the expected class rank for those who do not report a class rank puts this group at the 69th percentile. This number is comparable to the observable mean class rank for the data.

Table 5
First period performance regressions (1973 G.P.A.)

		One type		Two types	
		Coefficient	Stand. error	Coefficient	Stand. error
SAT math	Natural science	0.1502	0.0443	0.1156	0.0212
interactions	Business	0.1060	0.0504	0.0983	0.0328
(00's)	Soc/hum	0.0944	0.0330	0.1012	0.0170
	Education	0.0592	0.0431	0.0697	0.0280
SAT verbal	Natural science	0.0816	0.0425	0.1136	0.0260
interactions	Business	0.1100	0.0611	0.1170	0.0363
(00's)	Soc/hum	0.1590	0.0312	0.1510	0.0181
	Education	0.1053	0.0514	0.0935	0.0342
HS class rank	Natural science	1.0474	0.1972	1.0632	0.1921
interactions	Business	0.9004	0.2449	0.9065	0.2455
	Soc/hum	0.7607	0.1472	0.7283	0.1442
	Education	1.1111	0.2404	1.1516	0.2435
Female	Natural science	0.1586	0.0695	0.1441	0.0665
interactions	Business	0.1890	0.1128	0.2160	0.1105
	Soc/hum	0.0763	0.0525	0.0727	0.0520
	Education	0.1796	0.1040	0.1859	0.1005
Constant	Natural science	1.7394	0.2496	1.5799	0.2229
	Business	1.7607	0.2244	1.6681	0.2296
	Soc/hum	1.8962	0.1942	1.8608	0.1831
	Education	1.6176	0.2315	1.4863	0.2086
Coefficients	Do not know rank	0.6906	0.0499	0.6907	0.0484
common	Math quality[a]	−1.0544	0.6001	−1.7190	0.3133
across majors	Verbal quality	−0.4786	0.5594	0.0074	0.2788
Type 2	Natural science			0.3822	0.0593
interactions	Business			0.2360	0.0932
	Soc/hum			0.1406	0.0486
	Education			0.2749	0.0809

[a]The coefficients on college quality are constrained to be proportional to the major-specific coefficients on ability.

Without the mixture distribution, the coefficient on math college quality is negative one; a one point increase in both math ability and math college quality yields no change in expected grade point averages. The coefficient on verbal college quality is negative one-half, suggesting that grade inflation is more common at schools that have a disproportionately high verbal college quality. This result becomes magnified when the mixture distributions are added: schools with high math college qualities have grade deflation (the coefficient is less than negative one) with high verbal quality colleges having grade inflation.

Table 6
Second period performance regressions (1975 g.p.a.)

	One type		Two types	
	Coefficient	Stand. error	Coefficient	Stand. error
Constant	0.4532	0.0550	0.4022	0.0345
Expected g.p.a.[a]	0.8274	0.0347	0.8700	0.0207
First period shock	0.4459	0.0275	0.4400	0.0188
Shock × switch majors	−0.0757	0.0480	−0.0944	0.0327
Switch majors	−0.0297	0.0255	−0.0246	0.0180

[a]Calculated from the first period performance regression.

Females receive higher grades than their male counterparts. Larger effects are found in business with smaller, but still positive, effects in the social science/humanities. Adding unobserved heterogeneity has little effect on the female coefficients. The results with unobserved heterogeneity show that type 2's receive substantially higher grades in all subjects.

Table 6 displays the results of the second period performance regression. Adding the mixture distribution here only affects grade point averages through the predicted values from the first period regression. The expected grade point average was positive and slightly increased with the controls for unobserved heterogeneity. Both with and without the mixture distribution, the coefficient on the shock was positive while the corresponding coefficient on the shock times switching was negative. Hence, information is being conveyed in the first period shocks, a portion of which is major specific. The fixed cost of switching majors on g.p.a. was negative, but small and insignificant whether or not controls for unobserved heterogeneity were included.

Using the coefficient estimates, it is possible to back out the signal to noise ratio for those who stay in the same major and those who switch. For those who stay in the same major the estimated signal to noise ratio is 0.54 and 0.51 without and with unobserved heterogeneity respectively. These numbers decrease to 0.45 and 0.40 if the individual switched majors. That the numbers are smaller when unobserved heterogeneity is added makes sense: more about what the individual knows is in expected g.p.a. rather than in the shock, while the transitory portion is still present in the shock. Hence, without unobserved heterogeneity we would be overestimating the informational content of the shock.

4.2. Log earnings regressions

Estimates of the log earnings equations are given in Table 7. A key to later identifying the coefficient on earnings in the utility function is to have a variable which is only in the log earnings regression. As previously discussed, I use average state earnings for both workers who graduated from college and those who did not graduate from college as the exclusion restriction. The coefficient is positive and significant, though the magnitude does drop when unobserved heterogeneity is added.

Table 7
Log earnings regressions[a]

		One type		Two types	
		Coefficient	Stand. error	Coefficient	Stand. error
	State average earnings	0.5938	0.0779	0.2925	0.0344
SAT math	Natural science	0.0425	0.0361	0.0506	0.0159
interactions	Business	0.0198	0.0391	0.0217	0.0172
(00's)	Soc/hum	0.0165	0.0165	0.0203	0.0100
	Education	0.0000	0.0370	0.0000	0.0163
	No college	0.0304	0.0126	0.0310	0.0055
SAT verbal	Natural science	0.0000	0.0317	0.0000	0.0140
interactions	Business	0.0000	0.0391	0.0000	0.0173
(00's)	Soc/hum	0.0000	0.0224	0.0151	0.0099
	Education	0.0005	0.0376	0.0000	0.0165
	No college	0.0000	0.0131	0.0000	0.0058
Math	Natural science	0.0617	0.0948	0.0119	0.0422
school quality	Business	0.0032	0.1339	0.0015	0.0589
interactions	Soc/hum	0.0557	0.0926	0.0647	0.0407
	Education	0.0000	0.0370	0.0000	0.0586
Verbal	Natural science	0.0000	0.1087	0.0000	0.0481
school quality	Business	0.0010	0.1385	0.0000	0.0609
interactions	Soc/hum	0.0000	0.0960	0.0000	0.0422
	Education	0.0140	0.1403	0.0196	0.0618
Grades	Natural science	0.0617	0.0515	−0.0808	0.0233
	Business	0.2742	0.0547	0.1578	0.0245
	Soc/hum	0.1076	0.0402	0.0283	0.0179
	Education	0.1264	0.0603	0.1722	0.0269
Type 2	Natural science			0.5411	0.0223
interactions	Business			0.4379	0.0228
	Soc/hum			0.4468	0.0152
	Education			−0.3144	0.0229
	No college			0.4470	0.0089

[a] Sex and private school interacted with major and sex and college interacted with year dummies were also included along with major-specific constant terms.

The ability and school quality coefficients are all constrained to be greater than zero. Throughout, it is math ability and college quality which matter; the constraint on verbal ability and college quality almost always binds. The highest returns for math ability are seen for natural science majors, while math school quality is most important for social science/humanities majors. Controlling for selection using the mixture distribution lowers the return to college quality for natural science majors while keeping the other college quality coefficients close to the case without unobserved heterogeneity.

Table 8
Premiums for different majors[a]

		Natural science (%)	Business (%)	Soc sci/hum (%)	Education (%)
One type	Males	19.7	15.9	9.4	−1.2
	Females	15.0	24.4	13.0	5.2
	Change in premium				
	+100 SAT math	2.1	1.8	−0.4	−2.3
	+100 math quality	5.2	−2.7	4.5	−0.8
Two types	Males	22.1	14.4	5.4	−1.2
	Females	22.0	27.6	5.7	6.0
	Change in premium				
	+100 SAT math	1.0	0.6	−0.8	−1.9
	+100 math quality	2.8	−2.5	6.0	−2.1

[a]Premiums are relative to no college. Calculated using the average characteristics by sex for the sample.

Without the controls for unobserved heterogeneity, college grades are found to be an important contributor to future earnings. This is particularly the case for business majors; going from a 2.5 to a 3.0 yields an over 13% increase in yearly earnings. For the other majors, a similar increase in grade point average would yield around a 5% increase in earnings. With the exception of education majors, these effects diminish substantially when the mixture distribution is added. Now, going from a 2.5 to a 3.0 in business yields less than a 8% increase in yearly earnings. In fact, the coefficient on grades actually becomes negative for natural science majors. This may be due to an aggregation problem as biology majors may receive higher grades but lower earnings than the other natural science majors.

Types 2's received significantly higher earnings in all fields except for education. The type 2 coefficient for education is mitigated, however, by the positive effect type 2 has on grades in education. The implied correlations with grades suggest that the unobserved ability to perform well in school translates into higher earnings not only if the individual attended college, but also if the individual chose the no college option. Also included, but not reported, are private school interacted with field, sex interacted with field, year dummies interacted with college, and sex and year dummies interacted with college.

Whether premiums exist for particular majors is difficult to see given all of the interactions and the effect of ability and college quality through grades. Premiums for choosing one of the college majors over the no college option are displayed in Table 8 for both an average male and an average female. For the case with two types, I use the mean probabilities of being each of the types (0.5025 and 0.4875 for type 1's and type 2's, respectively). Also displayed are major-specific returns to math ability and college quality relative to the no college sector.[35] I do not analyze the effect of verbal

---

[35] These returns take into account the effect of math ability and college quality on earnings through grades.

ability and college quality as the constraint that the coefficients on these variables be greater than zero almost always binds.

Significant premiums for particular majors exist for both the average male and female ranging from a high of 27.6% (females in business, controlling for heterogeneity) to a low of −1.2% (males in education, whether or not we control for unobserved heterogeneity). The largest premiums are found in the natural science and business majors, implying that the gap in earnings across fields is not entirely driven by high ability individuals choosing the more lucrative fields. Adding unobserved heterogeneity had mixed effects on the premiums. Larger premiums existed for the natural science majors, but smaller premiums for the social science/humanities majors.

The total returns to math ability, both through grades and directly, are higher for natural science and business majors than in the no college sector. This is not the case for social science/humanities and education majors; from an earnings standpoint, increases in math ability make these two majors less attractive compared to the no college option. That is, while the returns to math ability are positive in the social sciences and education, they are actually higher in the no college sector. The returns to math college quality are positive for natural science and social science/humanities majors but negative for business and education majors. Even though the direct effects of college quality are constrained to be greater than zero, college quality can still have a negative effect through grades. This negative effect is stronger than the positive direct effect for business and education majors. The mixture distributions substantially lowered the returns to math college quality for natural science majors while increasing the returns for social science/humanities majors.

4.3. Estimates of the utility function

I next use the estimates of the performance and log earnings equations to obtain the second stage maximum likelihood estimates of the utility function parameters. Table 9 displays the maximum likelihood estimates for the parameters which are major specific. Sex and high school class rank interacted with major, along with major specific constant terms, also were included.

The first two sets of rows display the differences in preferences across abilities individuals have for each of the fields beyond the effort costs required in school. High math ability is more attractive for natural science and business majors, while high verbal ability is more attractive for social science/humanities majors. Controlling for unobserved heterogeneity had very little effect on these coefficients.

The effort costs, displayed in the next two rows, show that math ability is particularly useful in school. All of the math effort costs (in the form of relative math ability) are positive and significant, with a larger coefficient for natural science majors. While the magnitudes of these coefficients are reduced when the mixture distribution is added, they are still all positive and significant. On the other hand, the estimates reveal no significant verbal effort costs. Math ability, as in the log earnings regressions, seems to be much more important than verbal ability when predicting trends in major choice and returns to schooling. The one exception is that those with high verbal ability are attracted to the social science/humanities majors.

Table 9
Major-specific utility function parameters[a]

		One type		Two types	
		Coefficient	Stand. error	Coefficient	Stand. error
SAT math	Natural science	0.0511	0.0327	0.0482	0.0107
interactions	Business	0.0296	0.0214	0.0276	0.0095
(00's)	Soc/hum	0.0003	0.0072	0.0007	0.0065
	Education	0.0066	0.0102	0.0038	0.0088
SAT verbal	Natural science	−0.0051	0.0087	−0.0003	0.0078
interactions	Business	−0.0075	0.0108	−0.0052	0.0090
(00's)	Soc/hum	0.0320	0.0210	0.0315	0.0081
	Education	0.0018	0.0100	0.0039	0.0094
Relative	Natural science	0.2434	0.1002	0.1522	0.0443
math ability	Business	0.1118	0.0657	0.0702	0.0398
(00's)	Soc/hum	0.1232	0.0452	0.0861	0.0314
	Education	0.1155	0.0515	0.0801	0.0376
Relative	Natural science	0.0182	0.0385	0.0066	0.0306
verbal ability	Business	−0.0250	0.0459	−0.0105	0.0363
(00's)	Soc/hum	0.0646	0.0571	0.0283	0.0320
	Education	−0.0159	0.0436	−0.0147	0.0361
Performance	Natural science	3.3314	0.5818	3.2739	0.5287
shock	Business	1.9456	0.6320	1.7806	0.6018
	Soc/hum	3.0493	0.5626	2.6733	0.4846
	Education	2.2684	0.6228	1.8546	0.5935
Switching	Natural science	−2.0541	0.1878	−2.0075	0.1892
costs	Business	−2.6595	0.2039	−2.6802	0.2017
	Soc/hum	−0.8508	0.1377	−0.8307	0.1363
	Education	−2.5851	0.1984	−2.5692	0.1986
Type 2	Natural science			0.0808	0.0242
interactions	Business			0.0882	0.0195
	Soc/hum			0.1221	0.0178
	Education			0.1625	0.0164

[a]Sex and high school class rank interacted with major, along with major-specific constant terms, also were included.

Positive shocks to performance made staying in school more attractive, with stronger effects in the natural sciences and social science/humanities. The magnitudes of all the coefficients fall when the mixture distribution is added. This makes sense; previously one's type would be somewhat included in this performance shock. With an individual's type removed from the performance shock the information conveyed in the shock is not as relevant. Switching to a different major was very costly, though the costs were much smaller for switching into social science/humanities.

Table 10
Utility coefficients common across majors

	One type		Two types	
	Coefficient	Stand. error	Coefficient	Stand. error
Log earnings	2.2378	0.9775	0.8866	0.1450
Transfer schools	−2.1362	0.1348	−2.1630	0.1358
Private college	0.1415	0.0542	0.1022	0.0322
College in same state	−0.0323	0.0375	−0.0258	0.0282
Net cost[a] ($000's)	−0.0764	0.0287	−0.0526	0.170
Low income × net cost ($000's)	−0.1235	0.0333	−0.0889	0.0183
Relative ability squared	−0.0103	0.0057	−0.0070	0.0035
Discount factor	1.0081	0.0792	0.7312	0.0846
Differential preferences for[b]				
Jobs across abilities (ϕ_1)	5.1049	4.0973	8.6913	0.1529
College qualities (ϕ_2)	4.0071	2.4365	2.5600	0.7105
First period nesting parameter	0.6160	0.1057	0.6571	0.0865
Second period nesting parameter	0.6466	0.0979	0.6496	0.0789
Variance on first period decision[c]	0.3651	0.1332	0.2680	0.0965

[a]Calculated as tuition+books+room+board−scholarships.

[b]Preferences in the workplace are relative to the preferences in college (Section 3.4). See the first eight rows in Table 9.

[c]Variance on the second period decision is normalized to one.

Table 10 displays the utility coefficients that are common across majors. The coefficient on log earnings is positive and significant, though falls by more than half with the mixture distribution. Transferring schools is very costly, with a coefficient very similar to the coefficients on switching majors. The monetary cost of school acts as a deterrent to choosing a schooling option, with the effect stronger for those who come from a low income household. This suggests that low income households are either liquidity constrained or that their parents are paying a lower portion of their college expenses. Estimates of the yearly discount factor are 101% and 73% for the models without and with unobserved heterogeneity, respectively. Both coefficients are significantly different from zero.

Squared relative ability is negative and significant, suggesting that interior optimal school qualities may be a possibility. Further, much of the returns to college quality exist after the individual finishes college in the form of monetary returns and preferences for having attended a high-quality college. This suggests that the optimal first period college quality may be lower than the optimal second period college quality. Fixing verbal ability at the college verbal ability, I calculate the optimal gap between one's own math ability and that of the college if individuals were just maximizing the first period flow utility.[36] Without unobserved heterogeneity, the optimal math gaps ($A_M - \bar{A}_M$) are 19, −3, 61, and 45 for natural science, business, social science/humanities, and

---

[36] Included in this calculation is the fact that college quality serves as a consumption good as shown by the second row in the second set of Table 10.

education, respectively. This implies that first period flow utility is generally higher when individuals attend colleges with students who have lower math abilities than their own. These results are somewhat tempered when the mixture distribution is added with the optimal gaps now at 14, 0, 42, and 35.

Preferences for jobs associated with particular majors do differ across abilities as the estimate of ϕ_1 is much greater than one. Recall that if $\phi_1 = 0$ then there are no differences in utilities across abilities in the workplace except through earnings. Hence, the ability sorting that occurs across majors is in part driven by differences in preferences for particular jobs across abilities. With ϕ_2 being positive, college quality serves as a consumption good. Net of effort costs, individuals prefer to attend higher quality colleges. Recall that the reason there is just one number here was because of the restrictions placed on how preferences could vary across abilities, college qualities, and time. A likelihood ratio test that the restrictions hold cannot be rejected at the 90% level.

The nesting parameters are both relative to the variances of the no college error. These nesting parameters measure the cross-school component of the variance. In particular, had these coefficients been estimated to be one, then a multinomial logit would have resulted. That the actual estimates are less than one suggests that the preferences for schooling options are correlated.

5. Model fit

Given the parameter estimates, I now test how well the model matches the trends in the data. Table 11 compares the actual data on ability and college quality distributions to what the model predicts for both the first and second period choices. If the restrictions on the ability and college quality coefficients are wrong, then the estimated ability and college quality distributions will be wrong as well.[37] The data are indexed by 'D', with estimates without and with unobserved heterogeneity indexed by '1T' and '2T', respectively.

In all cases, the models with and without the mixture distributions predict the trends in the data very well. The models often hit the observed mean exactly. The only case where the distribution is slightly off is for business majors. For the first period choice, I overpredict by five and four points math ability and math college quality. Similarly, for the second period choice I underpredict math ability by six points. Both here and for the rest of this section adding unobserved heterogeneity did not improve the predictions.

While the model does a good job predicting the means, it is also important to see how well the model predicts the transitions. Table 12 displays both the actual and predicted transition matrix for math ability and for the percentage choosing each field. I focus on math ability because it has a much larger effect on both earnings and choice of major than verbal ability. The model predicts the percentage of people in each cell very well;

---

[37] I do not compare model predictions for the percentage of people choosing each major as these have to match in a nested logit framework when constant terms for each major and period are estimated.

Table 11
Model fit-abilities[a]

	Major		1972 choice		1974 choice	
			Own ability	Peer ability	Own ability	Peer ability
SAT math	Natural science	D	566	547	594	560
		1T	565	545	594	562
		2T	564	545	595	562
	Business	D	498	522	528	533
		1T	503	526	522	531
		2T	503	526	522	532
	Social science/ humanities	D	500	526	518	535
		1T	499	525	519	535
		2T	499	524	519	536
	Education	D	458	502	467	505
		1T	459	502	467	503
		2T	459	503	467	503
SAT verbal	Natural science	D	499	515	523	526
		1T	499	514	519	528
		2T	498	514	519	528
	Business	D	444	494	464	501
		1T	447	496	462	500
		2T	447	496	460	501
	Social science/ humanities	D	481	499	499	510
		1T	481	500	499	510
		2T	480	499	501	510
	Education	D	431	477	439	478
		1T	432	477	439	478
		2T	431	477	438	478

[a] 'D' refers to the actual data, '1T' to the estimates with one type, and '2T' to the estimates with two types.

including matching the higher drop-out rates found for the social science/humanities majors. Recall that it was much easier to switch into social science/humanities than the other majors. If an individual initially chose social science/humanities and did not like it, there is no low cost switch for him to make besides dropping out. While the model predictions often predict the observed distributions exactly, both models underpredict the drop-out rates of education majors by three percent.

Math ability transitions are also given. These do not match nearly as well, and this may in part be because of the small cell sizes. Looking along the diagonal, where most of the observations are, shows that the models predict the ability levels of those

Table 12
Model fit-transitions[a]

1972 Major	Variable		Natural science	1974 Business	Major Soc sci/ hum	Education	Drop out
Natural science	Own math ability	D	602	555	543	494	537
		1T	601	561	555	519	526
		2T	602	560	555	518	523
	% of '72 Major	D	42%	8%	19%	2%	28%
		1T	43%	6%	18%	4%	30%
		2T	43%	6%	18%	4%	30%
Business	Own math ability	D	579	516	506	498	455
		1T	558	516	510	475	474
		2T	562	516	511	475	472
	% of '72 Major	D	3%	54%	10%	3%	30%
		1T	3%	54%	12%	3%	29%
		2T	3%	54%	11%	3%	29%
	Own math ability	D	546	533	515	475	474
		1T	560	517	513	478	474
		2T	563	517	514	477	471
Soc sci/ humanities	% of '72 Major	D	4%	4%	49%	7%	36%
		1T	4%	5%	49%	6%	36%
		2T	4%	5%	49%	6%	36%
	Own math ability	D	583	537	479	460	433
		1T	537	494	491	457	443
		2T	541	495	492	457	441
Education	% of '72 Major	D	2%	3%	11%	51%	33%
		1T	2%	3%	12%	53%	30%
		2T	2%	3%	13%	52%	30%

[a] 'D' refers to the actual data, '1T' to the estimates with one type, and '2T' to the estimates with two types.

who stay in the same major well. The models also predict well the trend of decreasing math abilities along the rows. That is, higher math ability students are more likely to choose natural science over business, business over social science/humanities, and social science/humanities over education. However, the levels of the non-diagonal elements are off. Both models overpredict the abilities of those natural science majors who stay in school and, consequently, underpredict the abilities of natural science majors who drop out. In contrast, the model predicts average math ability of business dropouts to be 19 points higher than in the data.

Table 13
Simulations of the math ability distribution under different environments

| | | 1972 choice | | | | 1974 choice | | | |
| | | Simulations | | | | Simulations | | | |
		Base	(1)	(2)	(3)	Base	(1)	(2)	(3)
One type	Natural science	565	554	564	552	594	579	591	563
	Business	503	506	499	499	522	525	518	515
	Soc sci/hum	499	496	500	499	519	511	519	511
	Education	459	466	463	465	467	474	473	474
Two types	Natural science	564	554	563	549	595	582	592	551
	Business	503	504	500	497	522	522	518	509
	Soc sci/hum	499	496	496	498	519	512	515	509
	Education	459	466	458	467	466	474	465	478

Simulation (1): all colleges are of the same quality. Simulation (2): no monetary returns to ability or college quality. Simulation (3): no differences in job preferences based upon ability or college quality.

6. Simulation results

Given that the model matches the data reasonably well, I can use the model to simulate how the math ability distributions would vary given a different environment. The first simulation assumes that all individuals attend the same school. This simulation is designed to answer how much of the observed differences in math ability across majors is due to individuals attending schools of different quality levels. The second simulation turns off the returns to math and verbal ability as well as the returns to math and verbal college quality. The results of the simulation will then show how much of the ability sorting is due to differences in returns to abilities and college qualities. Finally, the third simulation turns off the differences in job preferences across abilities and college qualities ($\phi_1 = 0$). Note that these simulations are not taking into account general equilibrium effects; the simulations are only designed to illustrate how much of the current sorting is due to heterogeneous schools and returns to abilities.[38] Table 13 gives the results of the simulations as well as the estimates under the current environment.

The primary effect of having all schools be of uniform quality is to lower the average abilities of those choosing natural sciences while raising the average abilities of those who choose education. This occurs because higher quality colleges, which by definition are attended by high ability individuals, make the natural science majors more attractive. With everyone attending the same college, this effect is removed. Without unobserved heterogeneity, the gap between the average math abilities of second period natural science majors and the math abilities of second period business, social science/humanities, and education majors falls by 25%, 9%, and 17%, respectively.

---

[38] See Heckman et al. (1998) for what differences can arise between partial and general equilibrium policy simulations.

Falls of the same magnitude are present in the first period choice. With unobserved heterogeneity, the average math abilities do not fall as much. That is, part of the reason high-quality colleges have many natural science majors is due to selection. Hence, the gaps in math abilities between natural sciences and the other majors due to heterogeneous schools fall by 18%, 8%, and 16%, respectively.

While the effects are substantial for the policy simulation where everyone attends the same school, turning off the monetary returns to math ability and college quality has little effect on average math abilities across majors (simulation 2). The only effect of turning off the monetary returns is a small drop, two to four points, in average abilities for all second period majors when the mixture distribution is used. This occurs because the monetary returns to math ability are not substantially higher with a college degree and the monetary returns to college quality are small.

In contrast, removing the heterogeneity in preferences across abilities and college qualities for the jobs associated with each of the majors substantially reduces the ability sorting (simulation 3). For example, when the mixture distribution is used the gap between natural science majors and education majors falls by over 40% once heterogeneity in workplace preferences across abilities is removed. This still leaves the majority of the ability sorting occurring because of the school experience itself; individuals sort by ability in large part because of their preferences to study particular material while in school.

7. Conclusion

Large earnings and ability differences exist across majors. Selection into majors depends upon the monetary returns to various abilities, preferences in the workplace, and preferences for studying particular majors in college.

In order to separate out these components, I estimated a dynamic model of college and major choice. Individuals made an initial college and major decision conditional on expectations of what they would do in the future. After the initial choice, individuals received information about their preferences and, through their grades, about their abilities. With this new information, individuals updated their decisions by changing their major, changing their college, or entering the labor force. Estimates of the model revealed that positive ability shocks made staying in school attractive, especially for those interested in the natural sciences.

Math ability is found to be important both for labor market returns and for sorting into particular majors. In contrast, verbal ability has little effect on labor market outcomes or on sorting. Significant effort costs exist, with the effort being a function of the individual's math ability relative to his peers. These costs are convex and lead to interior optimal school qualities. While college quality serves as a consumption good, at the margin individuals trade off the costs of attending higher quality colleges with the benefits coming later in the form of both higher monetary returns and preferences for having attended a high-quality college.

Large monetary premiums exist for choosing natural science and business majors even after controlling for selection. However, these large premiums and the differential

monetary returns to ability and college quality cannot explain the ability sorting present across majors. Instead, virtually all sorting is occurring because of differing preferences across abilities for majors either in school or for the jobs associated with those majors in the workplace. While preferences for majors in school and preferences for the job associated with those majors in the workplace are both substantial, preferences in school play a larger role in the observed ability sorting.

Acknowledgements

References

Altonji, J., 1993. The demand for and return to education when education outcomes are uncertain. Journal of Labor Economics 11, 48–83.

Arcidiacono, P., Jones, J.B., 2003. Finite mixture distributions, sequential likelihood, and the EM algorithm. Econometrica 71, 933–946.

Berger, M., 1988. Predicted future earnings and choice of college major. Industrial and Labor Relations Review 41, 418–429.

Brewer, D., Ehrenberg, R., 1999. Does it pay to attend an elite private college? Cross cohort evidence on the effects of college quality on earnings. Journal of Human Resources 34, 104–123.

Cameron, S., Heckman, J., 1998. Life cycle schooling and dynamic selection bias: models and evidence for five cohorts of American males. Journal of Political Economy 106, 262–333.

Cameron, S., Heckman, J., 2001. The dynamics of educational attainment for black, hispanic, and white males. Journal of Political Economy 109, 455–499.

Card, D., 2001. Estimating the return to schooling: progress on some persistent econometric problems. Econometrica 69, 1127–1160.

Davidson, R., MacKinnon, J., 1993. Estimation and Inference in Econometrics. Oxford University Press, Oxford.

Daymont, T., Andrisani, P., 1984. Job preferences, college major, and the gender gap in earnings. Journal of Human Resources 19, 408–428.

Dempster, A.P., Laird, M., Rubin, D.B., 1977. Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistical Society, Series B 39, 1–38.

Eckstein, Z., Wolpin, K., 1989. The specification and estimation of dynamic stochastic discrete choice models: a survey. Journal of Human Resources 24, 562–598.

Eckstein, Z., Wolpin, K., 1999. Why youths drop out of high school: the impact of preferences, opportunities and abilities. Econometrica 67, 1295–1339.

Fuller, D., Manski, C., Wise, D., 1982. New evidence on the economic determinants of postsecondary schooling choices. Journal of Human Resources 17, 477–498.

Grogger, J., Eide, E., 1995. Changes in college skills and the rise in the college wage premium. Journal of Human Resources 30, 280–310.

Heckman, J., Vytlacil, E., 1998. Instrumental variables methods for the correlated random coefficient model. Journal of Human Resources 33, 974–987.

Heckman, J., Lochner, L., Taber, C., 1998. General-equilibrium treatment effects: a study of tuition policy. American Economic Review 88, 381–386.

Heckman, J., Tobias, J., Vytlacil, E., 2000. Simple estimators for treatment parameters in a latent variable framework with an application to estimating the returns to schooling. NBER Working Paper #7950.

James, E., Nabeel, A., Conaty, J., To, D., 1989. College quality and future earnings: where should you send your child to college?. American Economic Review 79, 247–252.

Keane, M., Wolpin, K., 1994. The solution and estimation of discrete choice dynamic programming models by simulation and interpolation: Monte Carlo evidence. Review of Economics and Statistics 76, 648–672.

Keane, M., Wolpin, K., 1997. The career decisions of young men. Journal of Political Economy 105, 473–522.

Loury, L., 1997. The gender-earnings gap among college-educated workers. Industrial and Labor Relations Review 50, 580–593.

Loury, L., Garman, D., 1995. College selectivity and earnings. Journal of Labor Economics 13, 289–308.

McFadden, D., 1981. Econometric models of probabilistic choice. In: Manski, C., McFadden, D. (Eds.), Structural Analysis of Discrete Data with Econometric Applications, MIT Press, Cambridge, MA.

Rothwell, G., Rust, J., 1997. On the optimal lifetime of nuclear power plants. Journal of Business and Economic Statistics 15, 195–208.

Rust, J., 1987. Optimal replacement of GMC bus engines: an empirical model of Harold Zurcher. Econometrica 55, 999–1033.

Rust, J., 1994. Structural estimation of Markov decision processes. Handbook of Econometrics, North-Holland, New York, Vol. 4.

Rust, J., 1996. Numerical dynamic programming in econometrics. Handbook of Computational Economics, Elsevier, New York.

Rust, J., Phelan, C., 1997. How social security and medicare affect retirement behavior in a world of incomplete markets. Econometrica 65, 781–831.

Sowell, T., 1972. Black Education: Myths and Tragedies. David McKay Company, Inc., New York.

Turner, S., Bowen, W., 1999. Choice of major: the changing (unchanging) gender gap. Industrial and Labor Relations Review 52, 289–313.

Willis, R., Rosen, S., 1979. Education and self selection. Journal of Political Economy 87, S7–S36.